JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **148**, 251–262 (1990)

The Least Extremal Solution of the Operator Equation AXB = C

PETR SLAVÍK

Department of Mathematics, Purkyně University (UJEP), Janáčkovo náměstí 2a, 662 95 Brno, Czechoslovakia

Submitted by G.-C. Rota

Received November 27, 1988

INTRODUCTION

In 1956, R. Penrose [4] proved that the least extremal solution of the matrix equation $\mathbf{AXB}=\mathbf{C}$ is a matrix $\mathbf{A}^{+}\mathbf{CB}^{+}$ (where $\mathbf{A}^{+}$ is the Moore–Penrose inverse). In so doing, he considered the so-called Frobenius matrix norm defined as $\|\mathbf{A}\|_2 := (\sum_{i,j} |a_{i,j}|^2)^{1/2}$. In 1980, Nashed and Engl [2] showed that the least extremal solution of the operator equation $\mathbf{AX}=\mathbf{C}$ with respect to a whole class of "Shatten" norms (including Hilbert–Schmidt norm) is an operator $\mathbf{A}^{+}\mathbf{C}$ (similarly also for the equation $\mathbf{XB}=\mathbf{C}$). The double-sided case, i.e., the equation $\mathbf{AXB}=\mathbf{C}$, has not been solved yet.

In this paper, Penrose's result is generalized to the operators in infinite-dimensional complex spaces for Hilbert–Schmidt norm. At first we show that the operator $\mathbf{A}^{\square}\mathbf{B}$ defined by $\mathbf{A}^{\square}\mathbf{B}: \mathbf{X} \mapsto \mathbf{AXB}$ has a generalized inverse and it satisfies $(\mathbf{A}^{\square}\mathbf{B})^{+}\mathbf{Y}=\mathbf{A}^{+}\mathbf{YB}^{+}$. To prove generalization of Penrose's result we employ the relation of $(\mathbf{A}^{\square}\mathbf{B})^{+}$ to generalized inverses $\mathbf{A}^{+}$ and $\mathbf{B}^{+}$ and the fact that we can introduce an inner product on the space $\mathbf{C}_2$.

We treat both bounded linear operators with closed range and arbitrary closed linear operators.

1. PRELIMINARIES

To fix the notation, throughout this paper let $\mathscr{B}, \mathscr{B}_1, \mathscr{B}_2, \ldots$ be the Banach spaces, $\mathscr{H}, \mathscr{H}_1, \mathscr{H}_2, \ldots$ be the Hilbert spaces (either all complex or all real), $\mathbf{L}(\mathscr{B}_1, \mathscr{B}_2)$ be the space of all bounded (and everywhere defined)

0022-247X/90 \$3.00

linear operators, and $\mathscr{C}(\mathscr{B}_1, \mathscr{B}_2)$ be the space of all densely defined closed linear operators from $\mathscr{B}_1$ to $\mathscr{B}_2$. For $\mathbf{L}(\mathscr{B}, \mathscr{B})$ we will write $\mathbf{L}(\mathscr{B})$. For a linear operator $\mathbf{T}$, we denote by $\mathscr{D}(\mathbf{T})$, $\mathscr{N}(\mathbf{T})$, $\mathscr{R}(\mathbf{T})$ its domain, nullspace, and range, respectively, $\mathbf{T}|_{\mathscr{M}}$ means the restriction of $\mathbf{T}$ to a subspace $\mathscr{M}$, and $\mathbf{P}_{\mathscr{M}}$ denotes the orthogonal projector onto the closed subspace $\mathscr{M}$. By $\mathscr{U} \dot{+} \mathscr{V}$ we denote the algebraic direct sum of two subspaces $\mathscr{U}$ and $\mathscr{V}$, by $\mathscr{W} = \mathscr{U} \oplus \mathscr{V}$ we denote the topological direct sum of two subspaces $\mathscr{U}$ and $\mathscr{V}$, i.e., $\mathscr{W} = \mathscr{U} \dot{+} \mathscr{V}$, and there exists a continuous projector $\mathbf{P}$ from $\mathscr{W}$ onto $\mathscr{U}$ such that $\mathscr{N}(\mathbf{P}) = \mathscr{V}$.

Let us first review some basic well-known results from the theory of generalized inverses of linear operators as they are presented in [3].

DEFINITION 1.1. Let $\mathscr{X}$, $\mathscr{Y}$ be topological vector spaces and $\mathbf{T}$ be a linear operator from $\mathscr{X}$ to $\mathscr{Y}$. Let $\mathbf{U}$ from $\mathscr{Y}$ to $\mathscr{X}$ be a densely defined linear operator with closed nullspace that fulfills the equations

$$\begin{aligned} \mathbf{TUT} &= \mathbf{T} && \text{on } \mathscr{D}(\mathbf{T}), \\ \mathbf{UTU} &= \mathbf{U} && \text{on } \mathscr{D}(\mathbf{U}), \\ \mathbf{TU} &= \mathbf{Q} && \text{on } \mathscr{D}(\mathbf{U}), \\ \mathbf{UT} &= \mathbf{I} - \mathbf{P} && \text{on } \mathscr{D}(\mathbf{T}), \end{aligned} \tag{1.1}$$

where $\mathbf{Q}$ is a (continuous) projector from $\mathscr{Y}$ onto $\overline{\mathscr{R}(\mathbf{T})}$, and $\mathbf{I} - \mathbf{P}$ is a projector from $\mathscr{X}$ onto $\overline{\mathscr{R}(\mathbf{U})}$. Then $\mathbf{U}$ is called a *topological generalized inverse* (abbreviated TGI) of $\mathbf{T}$ and we write $\mathbf{U} = \mathbf{T}^+ = \mathbf{T}^+_{\mathbf{P},\mathbf{Q}}$.

Remark. For our purposes we can assume that our toplogical spaces are Hausdorff.

DEFINITION 1.2. Let $\mathscr{D}(\mathbf{T}) \subseteq \mathscr{X}$. We say $\mathbf{T}$ is *domain decomposable with respect to the projector* $P \in L(\mathscr{X})$, if $\mathscr{N}(\mathbf{T}) \subseteq \mathscr{R}(\mathbf{P})$, $\mathbf{Px} \in \mathscr{N}(\mathbf{T})$ for all $\mathbf{x} \in \mathscr{D}(\mathbf{T})$, and $\mathscr{D}(\mathbf{T}) \cap \mathscr{N}(\mathbf{P})$ is dense in $\mathscr{N}(\mathbf{P})$.

THEOREM 1.3. *Let $\mathscr{X}$, $\mathscr{Y}$ be topological vector spaces, and* $\mathbf{T}$ *be a linear operator from $\mathscr{X}$ to $\mathscr{Y}$. If* $\mathbf{T}$ *is domain decomposable with respect to the (continuous) projector* $\mathbf{P}$ *and if there exists a (continuous) projector* $\mathbf{Q}$ *onto* $\overline{\mathscr{R}(\mathbf{T})}$ *then* $\mathbf{T}$ *has a unique TGI* $\mathbf{T}^+ = \mathbf{T}^+_{\mathbf{P},\mathbf{Q}}$ *(with respect to the choice of* $\mathbf{P}$ *and* $\mathbf{Q}$*).*

THEOREM 1.4. *Let $\mathscr{H}_1$, $\mathscr{H}_2$ be Hilbert spaces and* $\mathbf{T}$ *be a densely defined closed linear operator from $\mathscr{H}_1$ to $\mathscr{H}_2$. Then* $\mathbf{T}$ *has TGI. Special TGI* $\mathbf{T}^+$

corresponding to the choice of orthogonal projectors $\mathbf{P}$ *and* $\mathbf{Q}$ *is then equivalently characterized as a maximal linear operator* $\mathbf{X}$ *from* $\mathscr{H}_2$ *to* $\mathscr{H}_1$ *satisfying equations*

$$\begin{aligned} \mathbf{TXT} &= \mathbf{T} && \text{on} \quad \mathscr{D}(\mathbf{T}), \\ \mathbf{XTX} &= \mathbf{X} && \text{on} \quad \mathscr{D}(\mathbf{X}), \\ (\mathbf{TX})^* &= \mathbf{TX} && \text{on} \quad \mathscr{D}(\mathbf{X}), \\ (\mathbf{XT})^* &= \mathbf{XT} && \text{on} \quad \mathscr{D}(\mathbf{T}). \end{aligned} \tag{1.2}$$

Remark 1.5. Equations (1.2) are often used to a definition of generalized inverse in Hilbert spaces. It is denoted by $\mathbf{T}^+$ and called Moore–Penrose inverse or (orthogonal) generalized inverse.

DEFINITION 1.6. Let $\mathscr{N}_1, \mathscr{N}_2$ be normed spaces and $\mathbf{f}$ be a (generally nonlinear) mapping from $\mathscr{N}_1$ to $\mathscr{N}_2$. The vector $\hat{\mathbf{x}} \in \mathscr{D}(\mathbf{f})$ is called a *least extremal solution* or a *best approximate solution* of the equation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, if for all $\mathbf{x} \in \mathscr{D}(\mathbf{f})$ either

$$\|\mathbf{f}(\hat{\mathbf{x}}) - \mathbf{y}\| < \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|,$$

or

$$\|\mathbf{f}(\hat{\mathbf{x}}) - \mathbf{y}\| = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\| \qquad \text{and} \qquad \|\hat{\mathbf{x}}\| < \|\mathbf{x}\|.$$

THEOREM 1.7. *Let* $\mathbf{T} \in \mathscr{C}(\mathscr{H}_1, \mathscr{H}_2)$ *be a densely defined closed linear operator,* $\mathbf{T}^+$ *its (orthogonal) generalized inverse. Then for all* $\mathbf{y} \in \mathscr{D}(\mathbf{T}^+) = \mathscr{R}(\mathbf{T}) \oplus \mathscr{R}(\mathbf{T})^\perp$ *the vector* $\hat{\mathbf{x}} = \mathbf{T}^+\mathbf{y}$ *is the least extremal solution of the equation* $\mathbf{Tx} = \mathbf{y}$.

Remark. If we talk about mappings from the vector space to the other one, we always think of course of the spaces over the same field of scalars (either complex or real).

2. OPERATOR $\mathbf{A}^{\square}\mathbf{B}$ IN UNIFORM TOPOLOGY

Let $\mathscr{B}_1$, $\mathscr{B}_2$, $\mathscr{B}_3$, $\mathscr{B}_4$ be Banach spaces; $\mathbf{A} \in \mathscr{C}(\mathscr{B}_3, \mathscr{B}_4)$, $\mathbf{B} \in \mathscr{C}(\mathscr{B}_1, \mathscr{B}_2)$. Let us assume there exist (continuous) projectors $\mathbf{P}: \mathscr{B}_3 \to$ onto $\mathscr{N}(\mathbf{A})$, $\mathbf{Q}: \mathscr{B}_4 \to$ onto $\overline{\mathscr{R}(\mathbf{A})}$, $E: \mathscr{B}_1 \to$ onto $\mathscr{N}(\mathbf{B})$, and $\mathbf{F}: \mathscr{B}_2 \to$ onto $\overline{\mathscr{R}(\mathbf{B})}$, so that there exist topological generalized inverses $\mathbf{A}^+ = \mathbf{A}^+_{\mathbf{P},\mathbf{Q}}$, $\mathbf{B}^+ = \mathbf{B}^+_{\mathbf{E},\mathbf{F}}$. Thus we have the topological direct sum decompositions

$$\mathscr{B}_1 = \mathscr{N}(\mathbf{B}) \oplus \mathscr{N}(E) = \mathscr{N}(\mathbf{B}) \oplus \overline{\mathscr{R}(\mathbf{B}^+)};$$
$$\mathscr{B}_2 = \overline{\mathscr{R}(\mathbf{B})} \oplus \mathscr{N}(\mathbf{F}) = \overline{\mathscr{R}(\mathbf{B})} \oplus \mathscr{N}(\mathbf{B}^+);$$
$$\mathscr{B}_3 = \mathscr{N}(A) \oplus \mathscr{N}(\mathbf{P}) = \mathscr{N}(\mathbf{A}) \oplus \overline{\mathscr{R}(\mathbf{A}^+)};$$
$$\mathscr{B}_4 = \overline{\mathscr{R}(\mathbf{A})} \oplus \mathscr{N}(\mathbf{Q}) = \overline{\mathscr{R}(\mathbf{A})} \oplus \mathscr{N}(\mathbf{A}^+).$$

2.1. A, B-*Continuous Linear Operators with Closed Ranges*

Throughout this section, let **A** and **B** be continuous linear operators with closed ranges. Then $\mathbf{A}^+$ and $\mathbf{B}^+$ are everywhere defined bounded linear operators with closed ranges. These facts essentially simplify our considerations. We treat the case of arbitrary closed linear operators in the next section.

Let us now define an operator

$$\mathbf{A}^{\sqcup}\mathbf{B}\colon \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3) \to \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4) \qquad \text{by} \quad \mathbf{A}^{\sqcup}\mathbf{B}\mathbf{X} := \mathbf{A}\mathbf{X}\mathbf{B}. \tag{2.1}$$

Obviously, $\mathbf{A}^{\sqcup}\mathbf{B}$ is a linear continuous operator. It is now of interest to see if $\mathbf{A}^{\sqcup}\mathbf{B}$, as an operator between the Banach spaces $\mathbf{L}(\mathscr{B}_2, \mathscr{B}_3)$ and $\mathbf{L}(\mathscr{B}_1, \mathscr{B}_4)$, has a generalized inverse and, if so, to relate it to $\mathbf{A}^+$ and $\mathbf{B}^+$.

LEMMA 2.1. *For the nullspace and the range of* $\mathbf{A}^{\sqcup}\boldsymbol{B}$ *we have*

$$\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B}) = \{\mathbf{X} \in \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3) : \mathbf{A}\mathbf{X}\mathbf{B} = 0\},$$
$$\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B}) = \{\mathbf{Y} \in \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4) : \mathscr{R}(\mathbf{Y}) \subseteq \mathscr{R}(\mathbf{A}), \mathscr{N}(\mathbf{Y}) \supseteq \mathscr{N}(\mathbf{B})\}.$$

Proof. Representation of $\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})$ and the fact that $\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B}) \subseteq \{\mathbf{Y} \in \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4) : \mathscr{R}(\mathbf{Y} \subseteq \mathscr{R}(A), \mathscr{N}(Y) \supseteq \mathscr{N}(\mathbf{B})\}$ are obvious. Conversely, let $\mathbf{Y} \in \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4)$ be such that $\mathscr{R}(\mathbf{Y}) \subseteq \mathscr{R}(\mathbf{A})$ and $\mathscr{N}(\mathbf{Y}) \supseteq \mathscr{N}(\mathbf{B})$. Then $\mathbf{Y} = \mathbf{A}\mathbf{A}^+\mathbf{Y}\mathbf{B}^+\mathbf{B}$ and since $\mathbf{A}^+\mathbf{Y}\mathbf{B} \in \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3)$, we have $\mathbf{Y} = \mathbf{A}^{\sqcup}\mathbf{B}(\mathbf{A}^+\mathbf{Y}\mathbf{B}^+) \in \mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})$. ∎

LEMMA 2.2. *If*

$$\mathscr{M} := \{\mathbf{X} \in \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3) : \mathscr{R}(\mathbf{X}) \subseteq \mathscr{N}(\mathbf{P}) = \mathscr{R}(\mathbf{A}^+), \mathscr{N}(\mathbf{X}) \supseteq \mathscr{N}(\mathbf{E}) = \mathscr{N}(\mathbf{B}^+)\}$$

and

$$\mathscr{S} := \{\mathbf{Y} \in \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4) : \mathbf{A}^+\mathbf{Y}\mathbf{B}^+ = \mathbf{0}\},$$

then

$$\mathbf{L}(\mathscr{B}_2, \mathscr{B}_3) = \mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B}) \oplus \mathscr{M} \tag{2.2}$$

and

$$\mathbf{L}(\mathscr{B}_1,\mathscr{B}_4)=\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})\oplus\mathscr{S}. \tag{2.3}$$

Proof. The subspaces $\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})$, $\mathscr{M}$, $\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})$, and $\mathscr{S}$ are obviously closed. Let $\mathbf{X}\in\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})\cap\mathscr{M}$, i.e., $\mathbf{AXB}=\mathbf{0}$ and simultaneously $\mathscr{R}(\mathbf{X})\subseteq\mathscr{R}(\mathbf{A}^+)$, $\mathscr{N}(\mathbf{X})\supseteq\mathscr{N}(\mathbf{B}^+)$. Let $\mathbf{v}\in\mathscr{B}_2$ be arbitrary. We can write $\mathbf{v}=\mathbf{v}_1+\mathbf{v}_2$, where $\mathbf{v}_1\in\mathscr{N}(\mathbf{B}^+)$ and $\mathbf{v}_2\in\mathscr{R}(\mathbf{B})$, i.e., $\mathbf{v}_2=\mathbf{Bu}$ for some $\mathbf{u}\in\mathscr{B}_1$, Next $\mathscr{N}(\mathbf{B}^+)\subseteq\mathscr{N}(\mathbf{X})$, so that $\mathbf{Xv}_1=0$. Then $\mathbf{Xv}=\mathbf{Xv}_1+\mathbf{Xv}_2=\mathbf{XBu}$, but $\mathscr{R}(\mathbf{X})\subseteq\mathscr{R}(\mathbf{A}^+)$, hence $\mathbf{Xv}=\mathbf{A}^+\mathbf{AXBu}=0$. Thus $\mathbf{X}=\mathbf{0}$, hence $\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})\cap\mathscr{M}=\{\mathbf{0}\}$.

Similarly, let $\mathbf{Y}\in\mathscr{S}\cap\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})$, i.e., $\mathbf{A}^+\mathbf{YB}^+=\mathbf{0}$ and $\mathscr{R}(\mathbf{Y})\subseteq\mathscr{R}(\mathbf{A})$ and $\mathscr{N}(\mathbf{B})\subseteq\mathscr{N}(\mathbf{Y})$, which gives $\mathbf{Y}=0$, hence $\mathscr{S}\cap\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})=\{\mathbf{0}\}$.

Now, any $\mathbf{X}\in\mathbf{L}(\mathscr{B}_2,\mathscr{B}_3)$ can be written in the form $\mathbf{X}=\mathbf{A}^+\mathbf{AX}+(\mathbf{I}-\mathbf{A}^+\mathbf{A})\mathbf{X}=\mathbf{A}^+\mathbf{AXBB}^++\mathbf{A}^+\mathbf{AX}(\mathbf{I}-\mathbf{BB}^+)+(\mathbf{I}-\mathbf{A}^+\mathbf{A})\mathbf{X}=\mathbf{X}_1+\mathbf{X}_2$, where $\mathbf{X}_1=\mathbf{A}^+\mathbf{AXBB}^+$, $\mathbf{X}_2=\mathbf{A}^+\mathbf{AX}(\mathbf{I}-\mathbf{BB}^+)+(\mathbf{I}-\mathbf{A}^+\mathbf{A})\mathbf{X}$. Now obviously, $\mathscr{R}(\mathbf{X}_1)\subseteq\mathscr{R}(\mathbf{A}^+)$ and $\mathscr{N}(\mathbf{X}_1)\supseteq\mathscr{N}(\mathbf{B}^+)$ so that $\mathbf{X}_1\in\mathscr{M}$. Next $\mathbf{AX}_2\mathbf{B}=\mathbf{A}(\mathbf{A}^+\mathbf{AX}(\mathbf{I}-\mathbf{BB}^+)+(\mathbf{I}-\mathbf{A}^+\mathbf{A})\mathbf{X})\mathbf{B}=\mathbf{0}$, hence $\mathbf{X}_2\in\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})$. Thus (2.2) has been proved.

Similarly, any $\mathbf{Y}\in\mathbf{L}(\mathscr{B}_1,\mathscr{B}_4)$ can be written in the form $\mathbf{Y}=\mathbf{Y}_1+\mathbf{Y}_2$, where $\mathbf{Y}_1=\mathbf{AA}^+\mathbf{YB}^+\mathbf{B}$, $\mathbf{Y}_2=\mathbf{AA}^+\mathbf{Y}(\mathbf{I}-\mathbf{B}^+\mathbf{B})+(\mathbf{I}-\mathbf{AA}^+)\mathbf{Y}$. Then, obviously, $\mathbf{Y}_1\in\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})$, $\mathbf{Y}_2\in\mathscr{S}$, thus (2.3) has been proved. ∎

Let us denote $\mathfrak{P}:\mathbf{L}(\mathscr{B}_2,\mathscr{B}_3)\to\text{onto }\mathscr{N}(\mathbf{A}^{\sqcup}\mathbf{B})$ and $\mathfrak{Q}:\mathbf{L}(\mathscr{B}_1,\mathscr{B}_4)\to\text{onto }\mathscr{R}(\mathbf{A}^{\sqcup}\mathbf{B})$ the (continuous) projectors induced by (2.2) and (2.3), respectively. Then (from the proof of Lemma 2.2) we have

$$(\mathfrak{I}-\mathfrak{P})\mathbf{X}=(\mathbf{I}-\mathbf{P})\mathbf{XF},\qquad \mathfrak{Q}\mathbf{Y}=\mathbf{QY}(\mathbf{I}-\mathbf{E}).$$

THEOREM 2.3. *The operator* $\mathbf{A}^{\sqcup}\mathbf{B}$ *has the topological generalized inverse* $(\mathbf{A}^{\sqcup}\mathbf{B})^+=(\mathbf{A}^{\sqcup}\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}}$ *(with respect to the projectors* $\mathfrak{P}$ *and* $\mathfrak{Q}$*) and for all* $\mathbf{Y}\in\mathbf{L}(\mathscr{B}_1,\mathscr{B}_4)$ *we have*

$$(\mathbf{A}^{\sqcup}\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}}\mathbf{Y}=\mathbf{A}^+_{\mathbf{P},\mathbf{Q}}\mathbf{YB}^+_{\mathbf{E},\mathbf{F}}, \tag{2.4}$$

i.e.,

$$(\mathbf{A}^{\sqcup}\mathbf{B})^+=\mathbf{A}^{+\sqcup}\mathbf{B}^+.$$

Proof. It is seen from Lemma 2.2 that the operator $\mathbf{A}^{\sqcup}\mathbf{B}$ satisfies assumptions of Theorem 1.3, thus it has the TGI. It suffices to prove that an operator $\mathfrak{U}:\mathbf{L}(\mathscr{B}_1,\mathscr{B}_4)\to\mathbf{L}(\mathscr{B}_2,\mathscr{B}_3)$ defined as $\mathfrak{U}\mathbf{Y}:=\mathbf{A}^+\mathbf{YB}^+$ satisfies (1.3). The uniqueness of TGI then implies that $(\mathbf{A}^{\sqcup}\mathbf{B})^+=\mathfrak{U}$.

Let $\mathbf{X} \in \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3)$ and $\mathbf{Y} \in \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4)$ be arbitrary, but fixed. Then we have $\mathbf{A}^{\square}\mathbf{B}(\mathfrak{U}(\mathbf{A}^{\square}\mathbf{B}\mathbf{X})) = \mathbf{A}^{\square}\mathbf{B}(\mathfrak{U}(\mathbf{AXB})) = \mathbf{A}^{\square}\mathbf{B}(\mathbf{A}^{+}\mathbf{AXBB}^{+}) = \mathbf{AA}^{+}\mathbf{AXBB}^{+}\mathbf{B} = \mathbf{AXB} = \mathbf{A}^{\square}\mathbf{BX}$, similarly $\mathfrak{U}(\mathbf{A}^{\square}\mathbf{B}(\mathfrak{U}\mathbf{Y})) = \mathbf{A}^{+}\mathbf{AA}^{+}\mathbf{YB}^{+}\mathbf{BB}^{+}\mathbf{Y} = \mathbf{A}^{+}\mathbf{YB}^{+} = (\mathbf{A}^{\square}\mathbf{B})^{+}\mathbf{Y}$, next $\mathbf{A}^{\square}\mathbf{B}(\mathfrak{U}\mathbf{Y}) = \mathbf{AA}^{+}\mathbf{YB}^{+}\mathbf{B} = \mathbf{QY}(\mathbf{I}-\mathbf{E}) = \mathfrak{Q}\mathbf{Y}$, and finally $\mathfrak{U}(\mathbf{A}^{\square}\mathbf{BX}) = \mathbf{A}^{+}\mathbf{AXBB}^{+} = (\mathbf{I}-\mathbf{P})\mathbf{XF} = (\mathfrak{I}-\mathfrak{P})\mathbf{X}$, hence (2.4) is proved. ∎

Remark 2.4. It could be seen from Theorem 2.3 that the operations $^{\square}$ and $^{+}$ commute each other in a certain sense. We show in the next section that the commutativity holds also for more general operators.

2.2. A, B-*Arbitrary Closed Linear Operators*

Now we shall consider the case when $\mathbf{A} \in \mathscr{C}(\mathscr{B}_3, \mathscr{B}_4)$ and $\mathbf{B} \in \mathscr{C}(\mathscr{B}_1, \mathscr{B}_2)$ are arbitrary. To reach analogous results as in Section 2.1, we can proceed as follows:

We denote

$$\mathscr{D}_{\mathbf{Q}} := \mathscr{R}(\mathbf{A}) \oplus \mathscr{N}(\mathbf{Q}) \qquad (= \mathscr{D}(\mathbf{A}^{+})),$$
$$\mathscr{D}_{\mathbf{F}} := \mathscr{R}(\mathbf{B}) \oplus \mathscr{N}(\mathbf{F}) \qquad (= \mathscr{D}(\mathbf{B}^{+}))$$

the dense subspaces of $\mathscr{B}_4$ and $\mathscr{B}_2$, respectively, and

$$\mathscr{D}(\mathbf{A}, \mathbf{B}) := \{\mathbf{X} \in \mathbf{L}(\mathscr{D}_{\mathbf{F}}, \mathscr{D}(\mathbf{A})) : \mathbf{AXB} \in \mathbf{L}(\mathscr{D}(\mathbf{B}), \mathscr{B}_4)\},$$
$$\mathscr{Z}(\mathbf{A}, \mathbf{B}) := \{\mathbf{Y} \in \mathbf{L}(\mathscr{D}(\mathbf{B}), \mathscr{D}_{\mathbf{Q}}) : \mathbf{A}^{+}\mathbf{YB}^{+} \in \mathbf{L}(\mathscr{D}_{\mathbf{F}}, \mathscr{B}_3)\}$$

subspaces of the spaces $\mathbf{L}(\mathscr{D}_{\mathbf{F}}, \mathscr{B}_3)$ and $\mathbf{L}(\mathscr{D}(\mathbf{B}), \mathscr{B}_4)$, respectively.

We define now the operator

$$\mathbf{A}^{\square}\mathbf{B}: \mathscr{D}(\mathbf{A}, \mathbf{B}) \to \mathscr{Z}(\mathbf{A}, \mathbf{B})$$

by

$$\mathbf{A}^{\square}\mathbf{BX} := \mathbf{AXB}. \tag{2.5}$$

Obviously, $\mathbf{A}^{\square}\mathbf{B}$ is a correctly defined linear operator. Now we can proceed along the lines of Section 2.1.

Analogously as in Section 2.1, it is easy to verify that

$$\mathscr{N}(\mathbf{A}^{\square}\mathbf{B}) = \{\mathbf{X} \in \mathscr{D}(\mathbf{A}, \mathbf{B}) : \mathbf{AXB} = \mathbf{0}\},$$

and

$$\mathscr{R}(\mathbf{A}^{\square}\mathbf{B}) = \{\mathbf{Y} \in \mathscr{Z}(\mathbf{A}, \mathbf{B}) : \mathscr{R}(\mathbf{Y}) \subseteq \mathscr{R}(\mathbf{A}), \mathscr{N}(\mathbf{Y}) \supseteq \mathscr{N}(\mathbf{B})\}.$$

LEMMA 2.5. *If*

$$\mathscr{M} := \{\mathbf{X} \in \mathscr{D}(\mathbf{A}, \mathbf{B}) : \mathscr{R}(\mathbf{X}) \subseteq \mathscr{R}(\mathbf{A}^{+}), \mathscr{N}(\mathbf{X}) \supseteq \mathscr{N}(\mathbf{B}^{+})\}$$

and

$$\mathscr{S} := \{\mathbf{Y} \in \mathscr{Z}(\mathbf{A}, \mathbf{B}) : \mathbf{A}^+\mathbf{Y}\mathbf{B}^+ = \mathbf{0}\},$$

then

$$\mathscr{D}(\mathbf{A}, \mathbf{B}) = \mathscr{N}(\mathbf{A}^\square\mathbf{B}) \oplus \mathscr{M} \tag{2.6}$$

and

$$\mathscr{Z}(\mathbf{A}, \mathbf{B}) = \mathscr{R}(\mathbf{A}^\square\mathbf{B}) \oplus \mathscr{S}. \tag{2.7}$$

Proof. Analogously as in the proof of Lemma 2.2 it is easy to show that

$$\mathscr{D}(\mathbf{A}, \mathbf{B}) = \mathscr{N}(\mathbf{A}^\square\mathbf{B}) \dot{+} \mathscr{M} \tag{$*$}$$

and

$$\mathscr{Z}(\mathbf{A}, \mathbf{B}) = \mathscr{R}(\mathbf{A}^\square\mathbf{B}) \dot{+} \mathscr{S}. \tag{$**$}$$

Let $\mathfrak{P}$ be a projector from $\mathscr{D}(\mathbf{A}, \mathbf{B})$ onto $\mathscr{N}(\mathbf{A}^\square\mathbf{B})$ induced by ($*$), and $\mathfrak{Q}$ be a projector from $\mathscr{Z}(\mathbf{A}, \mathbf{B})$ onto $\mathscr{R}(\mathbf{A}^\square\mathbf{B})$ induced by ($**$). Then (similarly as in Section 2.1) the projectors $\mathfrak{P}$ and $\mathfrak{Q}$ satisfy

$$(\mathfrak{I} - \mathfrak{P})\mathbf{X} = \mathbf{A}^+\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{B}^+ = (\mathbf{I} - \mathbf{P})|_{\mathscr{D}(\mathbf{A})}\mathbf{X}\mathbf{F}|_{\mathscr{D}_\mathbf{F}}$$

and

$$\mathfrak{Q}\mathbf{Y} = \mathbf{A}\mathbf{A}^+\mathbf{T}\mathbf{B}^+\mathbf{B} = \mathbf{Q}|_{\mathscr{D}_\mathbf{Q}}\mathbf{Y}(\mathbf{I} - \mathbf{E})|_{\mathscr{D}(\mathbf{B})}.$$

Thus projectors $\mathfrak{P}$ and $\mathfrak{Q}$ are continuous, hence (2.6) and (2.7) have been proved. ∎

THEOREM 2.6. *The operator* $\mathbf{A}^\square\mathbf{B}$: $\mathscr{D}(\mathbf{A}, \mathbf{B}) \to \mathscr{Z}(\mathbf{A}, \mathbf{B})$ *defined by* (2.5) *has a topological generalized inverse*

$$(\mathbf{A}^\square\mathbf{B})^+ = (\mathbf{A}^\square\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}} : \mathscr{Z}(\mathbf{A}, \mathbf{B}) \to \mathscr{D}(\mathbf{A}, \mathbf{B})$$

and for all $\mathbf{Y} \in \mathscr{Z}(\mathbf{A}, \mathbf{B})$ *we have*

$$(\mathbf{A}^\square\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}}\mathbf{Y} = \mathbf{A}^+_{\mathbf{P},\mathbf{Q}}\mathbf{Y}\mathbf{B}^+_{\mathbf{E},\mathbf{F}}, \tag{2.8}$$

i.e.,

$$(\mathbf{A}^\square\mathbf{B})^+\mathbf{Y} = \mathbf{A}^+\mathbf{Y}\mathbf{B}^+.$$

Proof. It is easy to verify that the operator $\mathbf{A}^\square\mathbf{B}$ satisfies assumptions of Theorem 1.5, hence it has a TGI. Quite analogously as in proof of

Theorem 2.3 we could show that an operator $\mathfrak{U}: \mathscr{X}(\mathbf{A}, \mathbf{B}) \to \mathscr{D}(\mathbf{A}, \mathbf{B})$ defined by $\mathfrak{U}\mathbf{Y} := \mathbf{A}^+\mathbf{Y}\mathbf{B}^+$ satisfies (1.3). Then the uniqueness of TGI implies that $(\mathbf{A}^{\sqcup}\mathbf{B})^+ = \mathfrak{U}$. ∎

Remark 2.7. In the case of continuous operators **A**, **B** having closed ranges we have $\mathscr{D}_{\mathbf{Q}} = \mathscr{B}_4$, $\mathscr{D}_{\mathbf{F}} = \mathscr{B}_2$, $\mathscr{D}(\mathbf{A}, \mathbf{B}) = \mathbf{L}(\mathscr{B}_2, \mathscr{B}_3)$, $\mathscr{X}(\mathbf{A}, \mathbf{B}) = \mathbf{L}(\mathscr{B}_1, \mathscr{B}_4)$, so that Theorem 2.3 is a corollary of Theorem 2.6.

Remark 2.8. Let us define (analogously as the operator $\mathbf{A}^{\sqcup}\mathbf{B}$) an operator

$$\mathbf{A}^{+\sqcup}\mathbf{B}^+ : \mathscr{D}(\mathbf{A}^+, \mathbf{B}^+) \to \mathscr{X}(\mathbf{A}^+, \mathbf{B}^+)$$

by

$$\mathbf{A}^{+\sqcup}\mathbf{B}^+\mathbf{Y} := \mathbf{A}^+\mathbf{Y}\mathbf{B}^+,$$

where

$$\mathscr{D}(\mathbf{A}^+, \mathbf{B}^+) := \{\mathbf{Y} \in \mathbf{L}(\mathscr{D}(\mathbf{B}), \mathscr{D}(\mathbf{A}^+)) : \mathbf{A}^+\mathbf{Y}\mathbf{B}^+ \in \mathbf{L}(\mathscr{D}(\mathbf{B}^+), \mathscr{B}_3)\}$$

and

$$\mathscr{X}(\mathbf{A}^+, \mathbf{B}^+) := \{\mathbf{X} \in \mathbf{L}(\mathscr{D}(\mathbf{B}^+), \mathscr{D}(\mathbf{A})) : \mathbf{A}\mathbf{X}\mathbf{B} \in \mathbf{L}(\mathscr{D}(\mathbf{B}), \mathscr{B}_4)\}.$$

Obviously, $\mathscr{D}(\mathbf{A}^+, \mathbf{B}^+) = \mathscr{X}(\mathbf{A}, \mathbf{B})$, $\mathscr{X}(\mathbf{A}^+, \mathbf{B}^+) = \mathscr{D}(\mathbf{A}, \mathbf{B})$, hence we have

$$(\mathbf{A}^{\sqcup}\mathbf{B})^+ = \mathbf{A}^{+\sqcup}\mathbf{B}^+.$$

Thus the operations $^{\sqcup}$ and $^+$ commute in the sense mentioned above.

3. The Least Extremal Solution of the Equation $\mathbf{AXB} = \mathbf{C}$

In this section we show how to use results of Section 2 to generalize Penrose's assertion.

Throughout this section we shall consider the spaces over the complex field.

First, we review the definition of Hilber–Schmidt norm and some important properties of Hilbert–Schmidt operators as they are presented in [1].

Let $\mathbf{T} \in \mathbf{L}(\mathscr{H}_1, \mathscr{H}_2)$ be compact. Then the operator $\sqrt{\mathbf{T}^*\mathbf{T}}$ is correctly defined and it is compact, selfadjoint, and positive. Let $\{\mu_i(\mathbf{T})\}_{i=1}^{\infty}$, $\mu_1 \geqslant \mu_2 \geqslant \cdots$, be the sequence of eigenvalues of $\sqrt{\mathbf{T}^*\mathbf{T}}$, where each eigenvalue is repeated according to its multiplicity (if the sequence is finite, we complete it by zeroes). Now, we can define for $0 < p < \infty$ the functional $|\cdot|_p$ by

$$|\mathbf{T}|_p := \left(\sum_{n=1}^{\infty} (\mu_n)^p\right)^{1/p}.$$

We denote by $\mathbf{C}_p(\mathscr{H}_1, \mathscr{H}_2)$ the class of all compact operators $\mathbf{T}\in\mathbf{L}(\mathscr{H}_1, \mathscr{H}_2)$ for which $|\mathbf{T}|_p<\infty$. In [1] it is proved that for $1\leqslant p<\infty$, $\mathbf{C}_p(\mathscr{H}_1, \mathscr{H}_2)$ is a complete normed space with norm $|\cdot|_p$.

$|\cdot|_2$ is called Hilbert–Schmidt norm and operators from the class $\mathbf{C}_2$ are called Hilbert–Schmidt operators.

If $\mathbf{T}\in\mathbf{C}_p$ and $\mathbf{A}, \mathbf{B}$ are such bounded linear operators that operators $\mathbf{AT}$ and $\mathbf{TB}$ are defined, then $\mathbf{AT}$ and $\mathbf{TB}$ are also from the class $\mathbf{C}_p$ and we have $|\mathbf{AT}|_p\leqslant\|\mathbf{A}\|\cdot|\mathbf{T}|_p$ and $|\mathbf{TB}|_p\leqslant\|\mathbf{B}\|\cdot|\mathbf{T}|_p$. If $\mathbf{A}, \mathbf{B}\in\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_2)$, then $\mathbf{B}^*\in\mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_1)$ and $\mathbf{B}^*\mathbf{A}\in\mathbf{C}_1(\mathscr{H}_1, \mathscr{H}_1)$.

For each $\mathbf{T}\in\mathbf{C}_1(\mathscr{H}, \mathscr{H})$ we can define the functional $\operatorname{tr}(\mathbf{T}):=\sum_{i=1}^{\infty}\lambda_i(\mathbf{T})$, where $\lambda_i(\mathbf{T})$ are eigenvalues of $\mathbf{T}$. In [1] it is proved that tr is a continuous linear functional on $\mathbf{C}_1(\mathscr{H}, \mathscr{H})$.

By means of the foregoing considerations, it is easy to show that we can introduce the inner product on $\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_2)$ by the formula

$$\langle\mathbf{A}, \mathbf{B}\rangle:=\operatorname{tr}(\mathbf{B}^*\mathbf{A}). \tag{3.1}$$

In particular, $\langle\mathbf{A}, \mathbf{A}\rangle=\operatorname{tr}(\mathbf{A}^*\mathbf{A})=\sum_{i=1}^{\infty}\lambda_i(\mathbf{A}^*\mathbf{A})=\sum_{i=1}^{\infty}(\mu_i(\mathbf{A}))^2=|\mathbf{A}|_2^2$, where $\mu_i(\mathbf{A})$ are eigenvalues of $\sqrt{\mathbf{A}^*\mathbf{A}}$. Hence Hilbert–Schmidt norm $|\cdot|_2$ is induced by the inner product (3.1), so that $\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_2)$ is a Hilbert space.

We shall study the operator equation

$$\mathbf{AXB}=\mathbf{C}, \tag{3.2}$$

where $\mathbf{A}\in\mathscr{C}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B}\in\mathscr{C}(\mathscr{H}_1, \mathscr{H}_2)$ are closed linear operators. We denote by $\mathbf{A}^+=\mathbf{A}^+_{\mathbf{P},\mathbf{Q}}\in\mathscr{C}(\mathscr{H}_4, \mathscr{H}_3)$, $\mathbf{B}^+=\mathbf{B}^+_{\mathbf{E},\mathbf{F}}\in\mathscr{C}(\mathscr{H}_2, \mathscr{H}_1)$ generalized inverses of $\mathbf{A}$ and $\mathbf{B}$, with respect to orthogonal projectors $\mathbf{P}, \mathbf{Q}$ and $\mathbf{E}, \mathbf{F}$, respectively, where $\mathbf{P}=\mathbf{P}_{\mathscr{N}(\mathbf{A})}$, $\mathbf{Q}=\mathbf{P}_{\overline{\mathscr{R}(\mathbf{A})}}$, $\mathbf{E}=\mathbf{P}_{\mathscr{N}(\mathbf{B})}$, and $\mathbf{F}=\mathbf{P}_{\overline{\mathscr{R}(\mathbf{B})}}$. Next we denote $\mathscr{D}_\mathbf{F}:=\mathscr{R}(\mathbf{B})\oplus\mathscr{R}(\mathbf{B})^\perp$ and $\mathscr{D}_\mathbf{Q}:=\mathscr{R}(\mathbf{A}\oplus\mathscr{R}(\mathbf{A})^\perp$.

3.1. A, B-*Continuous Linear Operators with Closed Ranges*

At first we deal with the case when $\mathbf{A}\in\mathbf{L}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B}\in\mathbf{L}(\mathscr{H}_1, \mathscr{H}_2)$ are bounded linear operators with closed ranges. In order to find the least extremal solution of (3.2) with respect to Hilbert–Schmidt norm, we consider $\mathbf{X}$ and $\mathbf{C}$ to be Hilbert–Schmidt operators, i.e., $\mathbf{X}\in\mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3)$ and $\mathbf{C}\in\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$.

THEOREM 3.1. *Let* $\mathbf{A}\in\mathbf{L}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B}\in\mathbf{L}(\mathscr{H}_1, \mathscr{H}_2)$ *be operators with closed ranges,* $\mathbf{C}\in\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$, *and* $\mathbf{X}$ *is allowed to vary in* $\mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3)$. *Then the least extremal solution* (*with respect to Hilbert–Schmidt norm*) *of Eq.* (3.2) *is the operator* $\hat{\mathbf{X}}=\mathbf{A}^+\mathbf{C}\mathbf{B}^+$.

Proof. In the same way as in Section 2.1 we can show that the operator $\mathbf{A}^{\sqcup}\mathbf{B}\colon\mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3)\to\mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$ defined by $\mathbf{A}^{\sqcup}\mathbf{B}\mathbf{X}:=\mathbf{AXB}$ has the TGI

$(\mathbf{A}^{\square}\mathbf{B})^+ = (\mathbf{A}^{\square}\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}}$ (with respect to the projectors $\mathfrak{P}$ and $\mathfrak{Q}$, where $(\mathfrak{I}-\mathfrak{P})\mathbf{X} = (\mathbf{I}-\mathbf{P})\mathbf{XF}$, $\mathfrak{Q}\mathbf{Y} = \mathbf{QY}(\mathbf{I}-\mathbf{E})$) and it fulfills $(\mathbf{A}^{\square}\mathbf{B})^+\mathbf{Y} = \mathbf{A}^+\mathbf{YB}^+$ for all $\mathbf{Y} \in \mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$.

Now we prove that $\mathfrak{P}$ is the orthogonal projector from $\mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3)$ onto $\mathscr{N}(\mathbf{A}^{\square}\mathbf{B})$, equivalently that $\mathscr{R}(\mathfrak{P}) \perp \mathscr{N}(\mathfrak{P})$. We know

$$\mathscr{R}(\mathfrak{P}) = \mathscr{N}(\mathfrak{I}\text{-}\mathfrak{P}) = \{\mathbf{X} \in \mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3) : (\mathbf{I}-\mathbf{P})\mathbf{XF} = \mathbf{0}\},$$

$$\mathscr{N}(\mathfrak{P}) = \mathscr{R}(\mathfrak{I}-\mathfrak{P}) = \{\mathbf{X} \in \mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3) : (\mathbf{I}-\mathbf{P})\mathbf{XF} = \mathbf{X}\}.$$

Now, let $\mathbf{X} \in \mathscr{R}(\mathfrak{P})$, $\mathbf{Z} \in \mathscr{N}(\mathfrak{P})$ be arbitrary, but fixed. Then $\langle \mathbf{X}, \mathbf{Z} \rangle = \langle \mathbf{X}, (\mathbf{I}-\mathbf{P})\mathbf{ZF} \rangle = \operatorname{tr}(((\mathbf{I}-\mathbf{P})\mathbf{ZF})^*\mathbf{X}) = \operatorname{tr}(\mathbf{FZ}^*(\mathbf{I}-\mathbf{P})\mathbf{X})$. Next let $\mathbf{u}$ be the eigenvector and λ the corresponding eigenvalue of the operator $\mathbf{FZ}^*(\mathbf{I}-\mathbf{P})\mathbf{X}$. Since $\mathbf{F}^2 = \mathbf{F}$, we have $\mathbf{F}(\lambda\mathbf{u}) = \lambda\mathbf{u}$, hence either $\lambda = 0$ or $\mathbf{Fu} = \mathbf{u}$ so that $\mathbf{FZ}^*(\mathbf{I}-\mathbf{P})\mathbf{Xu} = \mathbf{FZ}^*(\mathbf{I}-\mathbf{P})\,\mathbf{XFu} = \mathbf{FZ}^*\mathbf{0u} = 0$, thus $\lambda = 0$ again. Hence $\langle \mathbf{X}, \mathbf{Z} \rangle = 0$ and $\mathfrak{P}$ is the orthogonal projector.

Quite analogously we can show that $\mathscr{R}(\mathfrak{Q}) \perp \mathscr{N}(\mathfrak{Q})$, hence $\mathfrak{Q}$ is the orthogonal projector, too.

Thus $(\mathbf{A}^{\square}\mathbf{B})^+$ is the orthogonal generalized inverse of $\mathbf{A}^{\square}\mathbf{B}$ and from Theorem 1.7 we have that the least extremal solution of the equation $\mathbf{A}^{\square}\mathbf{BX} = \mathbf{C}$ is $\hat{\mathbf{X}} = (\mathbf{A}^{\square}\mathbf{B})^+\mathbf{C}$, which was to be proved. ▮

Remark 3.2. Let $\mathscr{H}_1$, $\mathscr{H}_2$, $\mathscr{H}_3$, $\mathscr{H}_4$ be finite-dimensional Hilbert spaces. Then we can identify the operators **A**, **B**, **C**, and **X** with matrix representing them. It is easy to verify that Hilbert–Schmidt norm in finite dimension coincides with Frobenius matrix norm $\|\cdot\|_2$. Hence Penrose's result is a special case of Theorem 3.1.

3.2. A, B-*Arbitrary Closed Linear Operators*

At first we introduce an analogy of the functional $|\cdot|_p$ and of inner product on $\mathscr{C}_2$. Let $\mathbf{T}$, $\mathbf{S}$ be densely defined linear operators from $\mathscr{H}_1$ to $\mathscr{H}_2$. (If $\mathbf{T}$ is bounded, we denote by $\bar{\mathbf{T}}$ its unique extension to an everywhere defined bounded linear operator.) We define a functional $\|\cdot\|_p$ in the following way: if $\mathbf{T}$ is bounded and $\bar{\mathbf{T}} \in \mathbf{C}_p$, then $\|\mathbf{T}\|_p := |\bar{\mathbf{T}}|_p$, otherwise $\|\mathbf{T}\|_p := \infty$. Let $\mathbf{T}, \mathbf{S}$ be such that $\bar{\mathbf{T}}, \bar{\mathbf{S}} \in \mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_2)$. We define the functional $\{\cdot, \cdot\}$ by

$$\{\mathbf{T}, \mathbf{S}\} := \langle \bar{\mathbf{T}}, \bar{\mathbf{S}} \rangle. \tag{3.3}$$

Clearly, $\{\mathbf{T}, \mathbf{T}\} = \|\mathbf{T}\|_2^2$. Since the functional $\|\cdot\|_p$ fulfills the norm axioms on the special classes of operators, we will call it "p-norm."

In this section we shall assume that $\mathbf{A} \in \mathscr{C}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B} \in \mathscr{C}(\mathscr{H}_1, \mathscr{H}_2)$ are arbitrary closed linear operators.

Since the operator **AXB** is not defined on all $\mathscr{H}_1$, by solution of (3.2) we will mean such an operator $\mathbf{X}$ that $\mathbf{AXBx}=\mathbf{Cx}$ holds for all $\mathbf{x}\in\mathscr{D}(\mathbf{B})$. Thus we can deal with the equivalent equation

$$\mathbf{AXB}=\tilde{\mathbf{C}}, \tag{3.2)'$$

where $\tilde{\mathbf{C}}=\mathbf{C}|_{\mathscr{D}(\mathbf{B})}$. For reasons of uniqueness and correct definition of the operator $\mathbf{AXB}$ we consider only such $\mathbf{X}$ for which $\mathscr{D}(\mathbf{X})=\mathscr{D}(\mathbf{B}^+)=\mathscr{R}(\mathbf{B})\oplus\mathscr{R}(\mathbf{B})^\perp$ and $\mathscr{R}(\mathbf{X})\subseteq\mathscr{D}(\mathbf{A})$, respectively.

If we want to generalize results of the foregoing section, we have to alter the set over which he minimization takes place and to make accurate what the operator $\mathbf{C}$ can be. Thus we assume that the operator $\tilde{\mathbf{C}}$ is an element of the set

$$\mathscr{X}_2(\mathbf{A},\mathbf{B}):=\{\mathbf{Y}\in\mathbf{L}(\mathscr{D}(\mathbf{B}),\mathscr{D}_{\mathbf{Q}}):\|\mathbf{Y}\|_2<\infty,\ \|\mathbf{A}^+\mathbf{Y}\mathbf{B}^+\|_2<\infty\},$$

and $\mathbf{X}$ is allowed to vary in

$$\mathscr{D}_2(\mathbf{A},\mathbf{B}):=\{\mathbf{X}\in\mathbf{L}(\mathscr{D}_{\mathbf{F}},\mathscr{D}(\mathbf{A})):\|\mathbf{X}\|_2<\infty,\ \|\mathbf{AXB}\|_2<\infty\}.$$

These requirements are quite natural, and they guarantee that the operators $\mathbf{AXB}$ and $\mathbf{A}^+\tilde{\mathbf{C}}\mathbf{B}^+$ $(=\mathbf{A}^+\mathbf{C}\mathbf{B}^+)$ are densely defined and that "2-norm" of the operators $\mathbf{A}^+\tilde{\mathbf{C}}\mathbf{B}^+$ and $\mathbf{AXB}-\mathbf{C}$ is finite. Clearly, if $\mathbf{X}\in\mathbf{L}(\mathscr{D}_{\mathbf{F}},\mathscr{D}(\mathbf{A}))$ such that $\|\mathbf{X}\|_2<\infty$ and $\|\mathbf{AXB}\|_2=\infty$, then $\|\mathbf{AXB}-\mathbf{C}\|_2=\infty$, too, hence $\mathbf{X}$ is not the extremal solution of (3.2). In fact we look for the extremal solution on the set of $\mathbf{X}\in\mathbf{L}(\mathscr{D}_{\mathbf{F}},\mathscr{D}(\mathbf{A}))$ such that $\|\mathbf{X}\|_2<\infty$, but for simplicity we exclude from our considerations such $\mathbf{X}$, for which $\|\mathbf{AXB}\|_2=\infty$.

Obviously, $\mathscr{D}_2(\mathbf{A},\mathbf{B})$ and $\mathscr{X}_2(\mathbf{A},\mathbf{B})$ are linear subspaces of the spaces $\mathbf{L}(\mathscr{D}(\mathbf{B}),\mathscr{D}_{\mathbf{Q}})$ and $\mathbf{L}(\mathscr{D}_{\mathbf{F}},\mathscr{D}(\mathbf{A}))$, respectively, with the inner product $\{\cdot,\cdot\}$ defined by (3.3).

Analogously as in Section 2.2 it could be shown that the operator $\mathbf{A}^\square\mathbf{B}\colon\mathscr{D}_2(\mathbf{A},\mathbf{B})\to\mathscr{X}_2(\mathbf{A},\mathbf{B})$ defined by $\mathbf{A}^\square\mathbf{B}\mathbf{X}:=\mathbf{AXB}$ has the topological generalized inverse $(\mathbf{A}^\square\mathbf{B})^+=(\mathbf{A}^\square\mathbf{B})^+_{\mathfrak{P},\mathfrak{Q}}$ (with respect to the projectors $\mathfrak{P}$ and $\mathfrak{Q}$, where $(\mathfrak{I}-\mathfrak{P})\mathbf{X}=(\mathbf{I}-\mathbf{P})|_{\mathscr{D}(\mathbf{A})}\mathbf{X}\mathbf{F}|_{\mathscr{D}_{\mathbf{F}}}$, $\mathfrak{Q}\mathbf{Y}=\mathbf{Q}|_{\mathscr{D}_{\mathbf{Q}}}\mathbf{Y}(\mathbf{I}-\mathbf{E})|_{\mathscr{D}(\mathbf{B})}$) and for all $\mathbf{Y}\in\mathscr{X}_2(\mathbf{A},\mathbf{B})$ we have $(\mathbf{A}^\square\mathbf{B})^+\mathbf{Y}=\mathbf{A}^+\mathbf{Y}\mathbf{B}^+$.

Let us show that $\mathfrak{P},\mathfrak{Q}$ are orthogonal projectors (with respect to the inner product $\{\cdot,\cdot\}$ on $\mathscr{D}_2(\mathbf{A},\mathbf{B})$ and $\mathscr{X}_2(\mathbf{A},\mathbf{B})$, respectively). To prove it, let $\mathbf{X}\in\mathscr{R}(\mathfrak{P})$, $\mathbf{Z}\in\mathscr{N}(\mathbf{P})$ are arbitrary, but fixed. Then $\{\mathbf{X},\mathbf{Z}\}=\langle\bar{\mathbf{X}},\bar{\mathbf{Z}}\rangle$. But $(\mathbf{I}-\mathbf{P})\bar{\mathbf{X}}F=\mathbf{0}$ and $(\mathbf{I}-\mathbf{P})\bar{\mathbf{Z}}\mathbf{F}=\bar{\mathbf{Z}}$, and from the proof of Theorem 3.4 we obtain that $\{\mathbf{X},\mathbf{Z}\}=0$, henc $\mathfrak{P}$ is the orthogonal projector. Similarly it should be proved that $\mathfrak{Q}$ is the orthogonal projector, too. From Theorem 1.7 (the completeness of the spaces is not essential in its proof; moreover, we can of course complete the spaces $\mathscr{D}_2(\mathbf{A},\mathbf{B})$ and $\mathscr{X}_2(\mathbf{A},\mathbf{B})$), we immediately have:

THEOREM 3.3. *Let* $\mathbf{A} \in \mathscr{C}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B} \in \mathscr{C}(\mathscr{H}_1, \mathscr{H}_2)$ *be closed linear operators,* $\mathbf{C} \in \mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$ *be such that* $\tilde{\mathbf{C}} \in \mathscr{L}_2(\mathbf{A}, \mathbf{B})$, *and* $\mathbf{X}$ *varies in the set* $\mathscr{D}_2(\mathbf{A}, \mathbf{B})$. *Then the least extremal solution (with respect to "2-norm"* $\|\cdot\|_2$) *of the equation*

$$\mathbf{AXB} = \mathbf{C}$$

is

$$\hat{\mathbf{X}} = \mathbf{A}^+\mathbf{C}\mathbf{B}^+.$$

Remark 3.4. Clearly, if $\mathbf{A}$ and $\mathbf{B}$ are bounded linear operators with closed ranges, we have $\mathscr{L}_2(\mathbf{A}, \mathbf{B}) = \mathbf{C}_2(\mathscr{H}_1, \mathscr{H}_4)$ and $\mathscr{D}_2(\mathbf{A}, \mathbf{B}) = \mathbf{C}_2(\mathscr{H}_2, \mathscr{H}_3)$, thus Theorem 3.2 is the special case of Theorem 3.3.

Remark 3.5. For simple formulation we now denote the obvious induced operator norm $\|\cdot\|$ by $\|\cdot\|_\infty$. Then it is easy to show:

Let $\mathbf{A} \in \mathscr{C}(\mathscr{H}_3, \mathscr{H}_4)$, $\mathbf{B} \in \mathscr{C}(\mathscr{H}_1, \mathscr{H}_2)$ be closed linear operators, $\mathbf{C} \in \mathbf{L}(\mathscr{H}_1, \mathscr{H}_4)$ be a bounded linear operator, and let there exist some p, $1 \leqslant p \leqslant \infty$, such that the equation

$$\mathbf{AXB} = \mathbf{C} \tag{3.2}$$

has a solution $\mathbf{X}_0$, for which $\|\mathbf{X}_0\|_p < \infty$. Then the operator

$$\hat{\mathbf{X}} := \mathbf{A}^+\mathbf{C}\mathbf{B}^+$$

is the solution of (3.2) with minimal "norm" $\|\cdot\|_p$.

REFERENCES

1. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators," Part II, "Spectral Theory. Self Adjoint Operators in Hilbert Space," Interscience, New York, 1963; Russian translation, Mir, Moscow, 1966.
2. H. W. ENGL AND M. Z. NASHED, New extremal characterization of generalized inverses of linear operators, *J. Math. Anal. Appl.* **82** (1981), 566–586.
3. M. Z. NASHED AND G. F. VOTRUBA, A unified operator theory of generalized inverses, *in* "Generalized Inverses and Applications" (M. Z. Nashed, Ed.), pp. 1–109, Academic Press, San Diego, CA/New York, 1976.
4. R. PENROSE, On best approximate solution of linear matrix equations, *Proc. Cambridge Philos. Soc.* **52** (1956), 17–19.